Incomplete Proofs and Program Synthesis
Extended abstract

G. Mints
Stanford University

January 10, 2002

The theory of deductive program synthesis and verification relies on complete proofs of specifications. Such proofs are assumed to be found by an automated deduction program or constructed manually using a proof-checking system. This contradicts practice: even in mathematics most proofs are very far from being complete, and verification of programs usually checks only "principal" parts. Our goal here is to support this practice by some existing and new theory.

1 Classical First Order Logic

When specifications do not require inductive proofs, the main program synthesis tool is Herbrand's theorem. For existential formulas $\exists x R(x)$ with quantifier-free $R(x)$ there is a transformation of any first order proof $\pi : \exists x R(x)$ into a set of witnesses $t_1, \ldots, t_n$ such that $R(t_1) \vee \ldots \vee R(t_n)$. The whole proof π is needed in the standard formulation, while in fact only quantifier inferences are used, and the whole propositional part is redundant. Predicate inferences contain mathematically and algorithmically interesting part of the proof; propositional part is usually the most labor-consuming and often non-interesting part.

An exact formulation of the observation above uses ϵ-calculus. Instead of quantifiers it has terms $\epsilon x A(x)$, read "some x satisfying A(x)". The only non-propositional axioms are *critical formulas*

$$A(t) \to A(\epsilon x A(x)) \tag{1}$$

Quantifiers are defined by

$$(\exists x A(x))^* := A^*(\epsilon x A^*(x)), \tag{2}$$

using the relation "there exists x satisfying A(x) iff A is satisfied by some x".

Lemma 1.1 *(cf. [3]). The translation * transforms propositional rules into propositional rules, the rule of $\exists$-elimination into the rule of substitution and the rule of $\exists$-introduction into a critical formula (plus substitution and propositional inferences).*

After that the first ϵ-theorem (cf. [3]) allows us to find instances for existential $\exists x R(x)$ depending only on critical formulas Cr that tautologically imply $R(\epsilon x R(x))$.

2 Constructive Proofs

The main pragmatic reason for having constructive or intuitionistic proofs is a possibility to extract programs from proofs $\pi : \exists x A(x)$ without any restriction for $A(x)$. In this new logic one cannot completely ignore the propositional part of π: implications contribute significantly into the complexity of the eventual program. Most program extraction methods here are based on functional interpretation that are based on Brouwer-Heyting-Kolmogorov interpretation of constructive logical connectives. These interpretations differ in the amount of information they need. For example modified realizability mr, a functional interpretation introduced by G. Kreisel, ignores negative premises of implications:

$$x\ mr(\neg A \to B) \equiv \neg A \to x\ mr\ B$$

Another manifestation of the same phenomenon is Harrop's theorem.

Theorem 2.1 *For arbitrary A,B, if $\neg A \to \exists x B(x)$ is derivable (in intuitionistic first or higher order logic, intuitionistic first or higher order arithmetic etc.), then $\neg A \to B(t)$ for some t is derivable in the same theory.*

In fact $\neg A$ can be replaced by any $\vee, \exists$-free formula C. Proofs of such lemmas C, even of number-theoretic identities, to say nothing about Riemann Hypothesis or Fermat's Last Theorem, can be very complicated, but they can be skipped if we are interested only in the program.

3 Arithmetic; ϵ-substitution Method

In the case of classical arithmetic, the most venerable method of extracting witnesses from the proofs of purely existential formulas $\exists x R(x)$ with quantifier-free $R(x)$ is Hilbert's ϵ-substitution method, [3, 5]. It works after the *-translation (2) was applied to extract from a given proof π a finite system Cr of critical formulas. The method generates a sequence of *ϵ-substitutions*

$$S_0, S_1, \ldots \tag{3}$$

Each of S_i has a form

$$(\epsilon x_1 A_1, n_1), \ldots, (\epsilon x_1 A_k, n_k) \tag{4}$$

for natural numbers $n_1, \ldots, n_k$. The goal is to find an ϵ-substitution (4) *solving* given system Cr of critical formulas, that is making Cr true after a substitution $(\epsilon x_1 A_1/n_1, \ldots, \epsilon x_1 A_k/n_k)$ and computation. If an implication $Cr \to R(\epsilon x R(x))$

is derivable without use of critical formulas, such a solving substitution provides an n satisfying $R(n)$.

W. Ackermann proved termination of the ϵ-substitution method for arithmetic. His proof is not simple (G. Kreisel considered it to be a version of the priority method) and resists extension to stronger systems. A new approach proposed by the present author admits extension to stronger systems using infinitary proofs in ϵ-calculus [5]. Expansion of a linear sequence (3) into a two-dimensional infinitary proof h^∞ adds new intuitions and enables new geometrical constructions, but seems to prevent computational treatment. We describe below such a treatment using incomplete proofs.

Consider a new formal system $PA\epsilon*$ in the language of arithmetical ϵ-terms. Derivable objects are *sequents*

$$(\epsilon x_1 A_1, u_1), \ldots, (\epsilon x_1 A_k, u_k) \tag{5}$$

where e_i are closed arithmetical ϵ-terms containing no proper closed ϵ-subterms, and $u_i \in \{?, ?^0, +\} \cup \mathbb{N}$. A sequent contains at most one component of the form $(e, +)$. Such a component provides an incomplete information to be replaced later by a natural number, resulting in (e, n). A "proof" of a sequent $(e, +), \Theta$ can be simply an axiom $\mathsf{AxA}((e, +), \Theta)$, promising a proof of a suitable sequent $(e, n), \Theta$ for some $n \in \mathbb{N}$ in the future. In this sense proofs in $PA\epsilon*$ are incomplete. Exact definitions below use notation from [5].

Definition 1 *Two sequents* Σ *and* Θ *are* multiplicable *if* $\Theta \cup \Sigma$ *is a function after identification* $(e, ?^0) \mapsto= (e, ?)$ *and* $(e, +) \mapsto= (e, n)$ *if the latter is present. In this case we write* $\Theta * \Sigma$ *for* $\Theta \cup \Sigma$.

Axioms

$\mathsf{AxF}(\Theta)$	Θ is ci
$\mathsf{AxS}(\Theta)$	Θ is solving
$\mathsf{AxH}_{e,v}(\Theta)$	e is the H-term, v is the H-value of Θ
$\mathsf{AxA}(\Theta)$	Θ is an arbitrary sequent

Rules of inference

$$\frac{(e, ?^0), \Theta \quad (e, +), \Theta}{\Theta}\,\mathsf{Cut}_e \qquad \frac{(e, ?), \Theta \quad (e, +), \Theta}{\Theta}\,\mathsf{CutFr}_e$$

$$\frac{(e, ?^0), \Upsilon \quad (e, +), \Theta}{(e, ?), \Upsilon * \Theta}\,\mathsf{R}_e \qquad \frac{\Theta}{\Theta}\,\mathsf{E}_r \qquad \frac{\Theta}{\Theta}\,\mathsf{D}$$

$$\frac{(e, ?), \Theta}{\Theta}\,\mathsf{Fr}_e \qquad \frac{(e, v), \Theta_{\le rk(e)}}{(e, ?), \Theta}\,\mathsf{H}_{e,v} \qquad \frac{\Theta}{\Sigma * \Theta}\,\mathsf{W}_\Sigma$$

An ordinal assignment for a derivation h is defined with a path π for the end-sequent of h as an additional argument. In fact π is used only in R_e-case, and we omit it in all other cases.

Definition 2 *Let $\delta_\Theta = 1$, if $(e,+)$ occurs in Θ, and e is needed for computing truth-values of critical formulas or the next ϵ-substitution; $\delta_\Theta = 0$ otherwise.*

$$o(h) := \begin{cases} 1 & \textit{if } h \equiv \mathsf{AxX}(\Theta),\ X \neq A \\ N_{|\mathcal{CR}(\Theta)|_\Theta} + 1 - \delta_\Theta & \textit{if } h \equiv \mathsf{AxA}(\Theta),\ \Theta \textit{ computes } \mathrm{Cr} \\ \omega + N_{|\mathrm{Cr}|_\Theta} - \delta_\Theta & \textit{if } h \equiv \mathsf{AxA}(\Theta) \textit{ otherwise} \\ o(h_1) + \textit{length}(\pi) + 1 + o(h_0) & \textit{if } h = \mathsf{R}_e h_0 h_1 \\ \omega^{o(h_0)} & \textit{if } h = \mathsf{E}_r h_0 \pi \\ \max(o(h_0), o(h_1)) + 1 & \textit{if } h \equiv \mathsf{Cut}_e h_0 h_1, \mathsf{CutFr} h_0 h_1 \\ o(h_0) + 1 & \textit{if } h = \mathcal{T} h_0,\ \mathcal{T} \equiv \mathsf{Fr}, \mathsf{H}, \mathsf{D}, \mathsf{W} \end{cases}$$

Theorem 3.1 *Every sequence (4) generated by the ϵ-substitution method can be effectively transformed into a sequence of proofs $h_0 : S_0, h_1 : S_1, \ldots$ such that $o(h_0) > o(h_1) > \ldots$. Hence ϵ-substitution method terminates.*

Ordinal assignment $o(h)$ and the construction of proofs h_i use ideas from [4] and definitions from [1]. It is natural to expect that these constructions and proofs can be extended to all subsystems of analysis (second order arithmetic) admitting proof-theoretic ordinal analysis via cut-elimination. This will constitute a progress in problem stated by D. Hilbert in [2].

References

[1] W. Buchholz, Explaining Gentzen's Consistency Proof within Infinitary Proof Theory, in: Gottlob, G., Leitsch, A., Mundici, D. (eds.) Computational Logic and Proof Theory. 5th Kurt Gödel Colloquium, KGC'97 Lecture Notes in Computer Science, Vol.1289. Springer 1997

[2] D.Hilbert, Probleme der Grundlegung der Mathematik, Math.Ann. 1929, 102, 1-9

[3] D.Hilbert, P.Bernays, Grundlagen der Mathematik, Bd.2, Springer, 1970

[4] G. Mints, Finite Investigations of Transfinite Derivations, J.Soviet Mathematics, 1978, 10, 548-596 (Russian original 1975)

[5] G. Mints, S. Tupailo, W. Buchholz, Epsilon Substitution Method for Elementary Analysis, Archive for Math. Logic (1996), 35, 103-130